

AUTOMATIC MUSIC TRANSCRIPTION

Vojtěch Matoušek

Bachelor Degree Programme (3), FEEC BUT

E-mail: xmatou28@stud.feec.vutbr.cz

Supervised by: Petr Petyovský

E-mail: petyovsky@feec.vutbr.cz

Abstract: This article deals with the automatic music transcription into the music notation by short Fourier transform with logarithmic scale. The result of this work is application, which converts the music records, in the wave file, into the music notation. But the algorithm can work only with music records, where plays one tone at one time.

Keywords: automatic music transcription, short Fourier transform with logarithmic scale

1 ÚVOD

Transkripce, neboli přepis, je rozšířeným fenoménem a v hudbě má několik různých významů. Transkripce může být považována za úpravu či nové zpracování notového zápisu původní skladby nebo přepis starší notace do dnešního notového záznamu. Třetí význam, který je předmětem této práce, je přepis hudební nahrávky do notového písma [1]. Moderní algoritmy hudební analýzy jsou na vysoké úrovni, ale zatím žádný algoritmus nedokáže uspokojivě nahradit v přepisu hudební nahrávky do notového hudebně trénovaného člověka. Pro člověka bez hudebního sluchu a vzdělání je vytvoření notového záznamu značně obtížné až nemožné.

2 BLOKOVÉ SCHÉMA NAVRŽENÉHO PROGRAMU

Na začátku programu zvolím vstupní soubor skladby ve formátu wave, který obsahuje informace o analyzované skladbě. Ze vstupního souboru načtu vzorkovací frekvenci a jednotlivé vzorky skladby. Detekci jednotlivých not provádím přes Krátkodobou Fourierovou transformaci s logaritmickým frekvenčním rozlišením, které se věnuji v kapitole (3). Zpracování výsledků detekce not obnáší přiřazení jednotlivých frekvencí k pojmenovaným notám. Nakonec všechny výsledky uložím do textového souboru pro snadnější kontrolu výsledků.



Obrázek 1: Blokové schéma programu

3 ALGORITMUS PRO DETEKCI NOT

Jak už bylo řečeno v předchozí kapitole detekci not provádím přes Krátkodobou Fourierovou transformaci s logaritmickým frekvenčním rozlišením, která spočívá v tom, že budeme počítat pouze frekvence, které odpovídají půltónům hudební stupnice podle zadaného rozsahu. Pro dvanáct půltónů oktávy je vztah pro výpočet frekvence tónů následující [2]:

$$f_k = (\sqrt[12]{2})^k \times f_{min} \quad (1)$$

kde f_{min} je minimální analyzovaná frekvence skladby a $k = \{0, 1, \dots, k_{max}\}$, kde k_{max} vypočítám podle následujícího vztahu [2]:

$$k_{max} = \frac{12 \times \log\left(\frac{f_{max}}{f_{min}}\right)}{\log(2)} \quad (2)$$

kde f_{max} je maximální analyzovaná frekvence skladby.

Krátkodobá Fourierova transformace se aplikuje na vzorky analyzované skladby postupně po krátkých úsecích, které jsem si stanovil na 100 milisekund. Přepočet na počet vzorků N_i v časovém intervalu provádím dle rovnice (3):

$$N_i = f_s \times 0,1 \quad (3)$$

kde f_s je vzorkovací frekvence.

Algoritmus pro výpočet krátkodobé Fourierovy transformace s logaritmickým frekvenčním rozlišením vypadá následovně [3]:

$$X[N, k] = \sum_{n=N \times N_i}^{((N+1) \times N_i) - 1} x(n) e^{-j2\pi f_k \frac{n}{f_s}} \quad (4)$$

kde $x(n)$ je n -tý vzorek zpracovávaného signálu, f_k popisuje frekvenci k -tého tónu a f_s je vzorkovací frekvence. Index $k = \{0, 1, \dots, k_{max}\}$ indexuje jednotlivé půltóny tónů a index $N = \{0, 1, \dots, N_{max}\}$ [3], kde N_{max} se určí podle vztahu (5) a odpovídá maximálnímu počtu časových intervalů skladby přes které se počítá Fourierova transformace:

$$N_{max} = \frac{S}{N_i} \quad (5)$$

kde S je celkový počet vzorků ve skladbě a N_i je interval analyzovaných vzorků vypočtený ze vztahu (3).

Výsledek detekce not se ukládá do matice X o velikosti N_{max} řádků a k_{max} sloupců. Řádky matice X značí časový interval skladby a sloupce matice odpovídají jednotlivým zkoumaným frekvencím tónů [3]. Pro určení tónů ve skladbě vyberu z každého řádku maximální číslo. Index tohoto čísla odpovídá hranému tónu v konkrétním čase. Pokud maximální číslo bude blízké nule, tak v daný čas ve skladbě nehraje žádný tón a jedná se o pomlku.

Algoritmus provádí také jednoduchou detekci nástupu a délky noty. Když uvažuji monofonní skladby, což znamená, že se jedná o skladby, kdy v jeden čas hraje jen jeden tón, tak nástup noty detekuji při změně frekvence na danou notu. Délka noty odpovídá času, po který se frekvence nebude měnit. Algoritmus tedy nedokáže analyzovat dvě zahrané noty za sebou. Detekuje jen jeden dlouhý tón a bude fungovat jedině pro monofonní skladby.

4 DOSAŽENÉ VÝSLEDKY DETEKCE NOT

Pro demonstraci dosažených výsledků uvádím analýzu monofonní skladby Pop Goes the Weasel, která je dlouhá 11 sekund. Pro kontrolu, výsledky zapisuju do textového souboru. Výsledky detekce

not u výše zmíněné skladby jsou v tabulce (1). Z tabulky je patrné, že detekované noty se od originálních not skladby liší ve třech případech.

Správná nota	Detekovaná nota	Délka detekované noty [s]	Správná nota	Detekovaná nota	Délka detekované noty [s]
C ₄	C ₄	0,8	E ₄	E ₄	0,3
D ₄	D ₄	0,8	G ₄	G ₄	0,2
E ₄	E ₄	0,2	E ₄	E ₄	0,2
G ₄	G ₄	0,2	C ₄	C ₄	0,8
E ₄	E ₄	0,3	A ₄	A ₄	0,5
C ₄	C ₄	1,5	A ₄	Ais ₄	0,1
D ₄	D ₄	0,7	pomlka	pomlka	0,1
E ₄	E ₄	0,7	D ₄	D ₄	0,5
C ₄	H ₃	0,1	E ₄	F ₄	0,3
C ₄	C ₄	1,5	E ₄	E ₄	0,7
D ₄	D ₄	0,7	C ₄	C ₄	0,3

Tabulka 1: Porovnání detekovaných not s notami originální skladby

5 ZÁVĚR

Doposud navržený algoritmus je implementován ve vývojovém prostředí MATLAB a umí pracovat jen s monofonním zvukem, což znamená, že umí vyhodnocovat pouze jeden tón v daném čase. V dalších krocích se budu zabývat vylepšením algoritmu pro detekci not i pro jednodušší polyfonní skladby, kde v jeden časový okamžik může znít více tónů najednou. Jako výstup programu nabude textový soubor, ale výsledek detekce not program uloží do MIDI souboru.

REFERENCE

- [1] Šrámková, Kateřina.: Problematika transkripce v hudebních skladbách [online]. Brno: 2016 [cit.22.3.2018]. Dostupné z URL: <https://is.muni.cz/th/263724/pdf_d/disertace.pdf> Disertační práce: Masarykova univerzita, Pedagogická fakulta. Vedoucí práce Michal Košut.
- [2] Brown, Judith C.: Calculation of a constant Q spectral transform [online]. 1990 Wellesley, Massachusetts [cit.22.3.2018]. Dostupné z URL: <<https://www.ee.columbia.edu/~dpwe/papers/Brown91-cqt.pdf>>
- [3] Schörkhuber, Ch. & Klapuri, A.: Constant-Q transform toolbox for music processing. Proc. 7th Sound and Music Computing Conference 2010.